



Gao, Z., Yan, S., Zhang, J., Mascarenhas, M., Nejabati, R., Ji, Y., & Simeonidou, D. (2020). ANN-Based Multi-Channel QoT-Prediction over a 563.4-Km Field-Trial Testbed. *Journal of Lightwave Technology*, 38(9), 2646-2655.
<https://doi.org/10.1109/JLT.2020.2971104>

Peer reviewed version

Link to published version (if available):
[10.1109/JLT.2020.2971104](https://doi.org/10.1109/JLT.2020.2971104)

[Link to publication record in Explore Bristol Research](#)
PDF-document

This is the author accepted manuscript (AAM). The final published version (version of record) is available online via Institute of Electrical and Electronics Engineers at <https://ieeexplore.ieee.org/document/8978723> . Please refer to any applicable terms of use of the publisher.

University of Bristol - Explore Bristol Research

General rights

This document is made available in accordance with publisher policies. Please cite only the published version using the reference above. Full terms of use are available:
<http://www.bristol.ac.uk/red/research-policy/pure/user-guides/ebr-terms/>

ANN-based Multi-Channel QoT-Prediction over a 563.4-km Field-Trial Testbed

Zhengguang Gao, *Student Member, IEEE*, Shuangyi Yan, Jiawei Zhang, Marcus Mascarenhas, Reza Nejabati, Yuefeng Ji, *Senior Member, IEEE*, and Dimitra Simeonidou, *Fellow, IEEE*

Abstract—In this paper, artificial neural network (ANN)-based multi-channel Q-factor prediction is investigated with real-time network operation and configuration information over a 563.4-km field-trial testbed. A unified ANN-based regression model is proposed and implemented to predict Q-factors of all the channels simultaneously. A scenario generator is developed to configure the field-trial testbed with 8 testing channels automatically to generate dynamic scenarios. A network configuration and monitoring database (CMDB) is implemented to collect network configuration and monitoring data that include link information, operational parameters of key optical devices, network configuration state, and real-time Q-factors of the available channels for the generated network scenarios. These collected data are used for training and testing of the developed ANN model. In order to achieve multiple channel predictions, we propose a hot coding method to represent the state of dynamic channel. Besides, an auto-search method is used to search the best hyperparameters of the ANN-based model. The results show that the proposed ANN-based regression model converges quickly, and it can predict the multi-channel's Q-factors with high accuracy. The unified ANN-based multi-channel Q-factor regression model can provide the comprehensive information to assist SDN controller to optimize network configuration for dynamic optical networks.

Index Terms—artificial neural network, multi-channel Q-factor prediction, field-trial testbed.

I. INTRODUCTION

THE emerging 5G applications will rise huge challenges for optical networks that underpin the digital communication infrastructure [2]–[3]. On the one hand, high capacity applications, such as high-resolution video and AR/VR, have been driving extreme traffic growths in optical networks, which requires further developments of high-capacity transmission systems and improvements on network utilization efficiency [4]. On the other hand, the increasing ratio of mobile-related traffic drives optical networks to be more dynamic in terms of various service times and on-demand bandwidths [5]. Both high-capacity requests and heterogenous user traffic require a dynamic optical network with automatic network

reconfigurations. Recent innovations in software defined networking (SDN) and network function virtualization (NFV) push optical networks evolving to be automatic and flexible in perspective of network functions and the control plane of optical networks [6–7]. SDN decouples the data plane and the control plane of optical networks, which enables independent developments of both the control and data plane. The centralized network control of SDN allows optical network to be programmable with flexibility. NFV technologies remove the restriction on network functions with dedicated hardware devices. On-demand network services can be synthesized on the general computing hardware through virtualization techniques. These flexibilities in network functions and managements will further bring more dynamic network traffic. It is expected in a foreseeable future that optical networks need to be reconfigured at the level of optical channels to handle the huge traffic variations.

By adapting optical networks in network functions and link bandwidth dynamically to the real-time network traffic, optical networks can improve resource utilization significantly for the short-term bursts of network traffic. However, dynamically configurable optical networks rise new changes for network operation and managements. To provide the user-defined services, optical networks are configured at short time scales to allocate resources efficiently. The on-demand network reconfigurations naturally require that low-margins are reserved to ensure high resource utilization as life cycles of the service reduce significantly [8]. Besides, the traditionally manual link configuration and optimization should be upgraded to support high-frequency network reconfigurations. Automated and cognitive network management strategies should be developed to organize network resources effectively and quickly [9].

To achieve automated network operations with low link margins, several challenges should be solved. Firstly, the centralized network controller needs precise real-time global view of network states. The low-margin design raises concern

Parts of this work appeared in the proceeding of ECOC 2019, Dublin, Ireland [1].

This work was supported by the National Key R&D Program of China (No. 2018YFB1800802), the National Nature Science Foundation of China Projects (61971055), the Beijing Natural Science Foundation (No. 4192039), the fund of State Key Laboratory of Advanced Optical Communication Systems and Networks, China, 2019GZKF5. The UK EPSRC through the TOUCAN project (EP/L020009/1) and the NDFIS project ((NS/A000021/1)). (*Corresponding author: Yuefeng Ji.*)

Zhengguang Gao, Jiawei Zhang, and Yuefeng Ji are with the State Key Laboratory of Information Photonics and Optical Communications, Beijing University of Posts and Telecommunications, 100876, China (email: gaozg@bupt.edu.cn, zjw@bupt.edu.cn, jyf@bupt.edu.cn).

Zhengguang Gao, Shuangyi Yan, Marcus Mascarenhas, Reza Nejabati, and Dimitra Simeonidou are with the High Performance Networks Group, University of Bristol, Bristol, U.K. ({cd18966, shuangyi.yan, mml15722.2015 reza.nejabati, dimitra.simeonidou}@bristol.ac.uk).

about network reliability. Therefore, each new connection should be maximized on total link capacities by adopting advanced-modulation signals to take full advantages of the obtained extra margin [10]. On the other side, the establishing connection should affect current connections in a manageable level and avoid possible failures of the current links. Precise predictions of Quality of Transmission (QoT) for both the unestablished channels and the current channels after adding new channels should be developed to assist the automatic low-margin network configurations.

Recently, QoT predications have become a hot topic. Different approaches have been explored. Traditional analytical models, such as Gaussian Noise model, have been simplified to achieve a fast QoT estimation. However, the simplified model requires extra resource margins to compensate for the uncertainty in the worst scenarios [11]. Sophisticated analytical models can provide accurate QoT prediction, but the fact that sophisticated model requires high computational resources makes it unsuitable for the real-time estimation in practical scenarios [12]. On the other side, the rapid developments of Machine Learning (ML) technologies attract many researchers to explore ML-based QoT estimations. Various ML-based methods have been proposed to predict QoTs of optical channels. ML-based QoT classification was proposed to predict whether the QoTs of channels exceed the pre-defined QoT thresholds without providing real values for predictions; ML-based QoT regression was investigated to predict the exact value of Optical Signal Noise Ratio (OSNR) or Q-factor, which can provide more details than the classification [13]. Some progress has been made in ML-based QoT estimations, various Machine Learning (ML) algorithms such as Support Vector Machine, Ensemble Learning methods, and artificial neural network (ANN) have been applied in QoT prediction [14-29]. However, there are several issues that prohibit the deployment of ML-based models for practical networks. Firstly, many works used synthetic data [14-17], [20-21], [24], [27], which are questionable about the effectiveness of the trained ML models for practical applications. To develop an effective and precise ML model, practical data from commercial optical networks are preferred to reflect both uncertainties and unknown factors. On the other hand, most existed works mainly focus on QoT prediction for a single channel, which cannot provide complete information for SDN controllers to make decisions.

In this paper, a unified ANN-based multi-channel Q-factor prediction model is proposed and implemented over a 563.4-km field-trial testbed, the UK National Dark Fiber Facility (NDFF), to predict multiple optical channels based on practical physical parameters. The collected data includes link configuration information and the operation parameters of the physical layer devices, such as EDFAs, optical switches. We extended our previous work in [1]. The ANN-based Q-factor regression model is redesigned to improve the generalization of the QoT prediction. And we extend the tests which consider new scenarios with 4 and 5 available channels. The influence of training data's size on the ANN-based model is also analyzed in terms of the accuracy. The results show that the developed

ANN-based model converges quickly. The learning curve of the ANN-model proves that the model achieves good generalization for the training data as well as the testing data. The unified ANN-based model can predict the Q-factors of the unestablished channels (Plan) and the existed channels together accurately in a real scenario. In addition, the unified ANN-based QoT prediction model makes the further scaling up for complex network scenarios, such as to predict a channel that passes several sections of fiber links with different channel loading. This precise multi-channel Q-factor prediction can assist SDN controller to optimize network configurations based on reduced link margins. The work paves a path for future self-organized autonomous networks.

The remainder of the paper is organized as follows. Some related works are presented and discussed in Section II. The field-trial testbed is described in Section III. The algorithm of ANN-based multi-channel Q-factor regression is proposed in Section IV. In Section V, we discuss the results in detail. The future work is presented in VI. At last, we make a conclusion in Section VII.

II. RELATED WORKS

Machine learning-based QoT estimation has been investigated comprehensively in recent years. Firstly, ML-based QoT classifier has been proposed to predict whether the performance of lightpaths meets the required QoT threshold. Reference [14] used a Random Forest-based and K-Nearest Neighbor (KNN)-based classifier to predict the QoT of unestablished lightpaths. The proposed classification model achieved a promising performance on synthetic data considering traffic volume, desired route, and modulation format. Reference [15] proposed an ML-based cognitive tool for QoT estimation for new lightpaths based on Random Forest, Support Vector Machine, and KNN. The performances of several ML-based classification algorithms were evaluated in three network scenarios [16]. The results showed that ANN achieved the best generalization with the accuracy of about 99%. The author in [17] proposed a graph convolutional neural network-based method to make classification for all the unseen lightpaths. The QoT classifier can provide some prior information to identify available lightpaths with satisfactory performance, however, it cannot give more detailed information to pick the best lightpath. Therefore, applications of ML for OSNR or Q-factor regression were investigated in the following works.

The ANN-based OSNR prediction were evaluated over a field-trial testbed [18]. The results showed that ANN-based regression model achieved a high accuracy that the mean square error (MSE) is less than 1 dB. The author in [19] experimentally demonstrated an ML-based QoT estimation over multi-domain elastic optical networks both for classification and regression tasks. The ANN-based estimator can predict the OSNR of a given channel within a 6% estimation error. In addition to building ML-based models directly, people also tried to improve the accuracy of existed physical analytical models based on Machine Learning. Reference [20] tried to reduce the uncertainty of parameters for

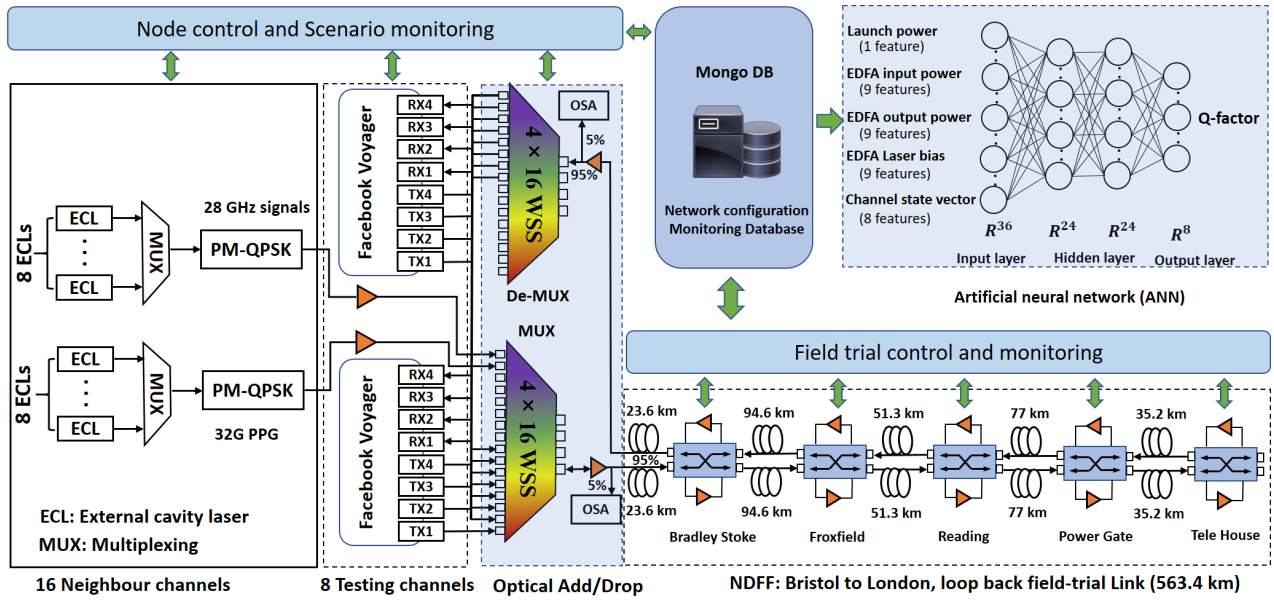


Fig. 1. 563.4-km field-trail testbed with cloud network configuration and monitoring database. Inset: the proposed ANN-based multi-channel QoT estimation.

a mathematical model to predict lightpath's SNR. The results showed that the learning process can reduce QoT prediction error significantly from 1.8 dB to 0.1 dB for the semi-analytical model and from 4.2 dB to 0.02 dB for the extended Gaussian noise model. The authors in [21] used two concepts of ML-assisted methods to predict the OSNR of lightpaths. The first method is to learn the parameters of the physical layer model for QoT estimation. The second is to predict OSNR directly by Machine Learning. The result showed that the maximum overestimation and minimum underestimation error for ML-based physical analytical model were about within 0.1 dB and 0.3 dB. For the direct ML-based method, the errors were about within 0.3 dB and 0.9 dB. An accurate ML-based model for QoT prediction usually requires large amounts of data for the training. Considering the time cost for training and the expense for collecting the data, the researchers began to explore transfer learning-based QoT estimations with small samples.

Reference [22] proposed an ANN-based transfer learning model for Q-factor estimation in optical networks. The results showed that transfer learning can help to reduce the number of training datasets effectively. A DNN-assisted transfer learning was proposed to make real-time OSNR monitoring [23], the results showed that transfer learning-aid DNN model can achieve fast remodeling for OSNR estimation without any accuracy decreasing. The authors in [24] discussed active learning to reduce the dataset needed for the training of ML-based QoT estimation. At least 5% and up to 75% of training samples is reduced by active learning to obtain the same results of traditional offline ML methods.

The accurate QoT estimation can assist the centralized controller to optimize network configuration, so the integration of ML-based QoT estimation and network planning had been investigated in advance. The authors in [25] firstly demonstrated an SDN-based optical network deploying ML to optimize network operation. The work tried to maximize the

link capacity through monitoring network state based on the ML-based OSNR prediction. Reference [26] demonstrated a 6 CD-ROADMs network testbed with ML-aid modulation format provisioning. About 15 percent of capacity increase had been achieved through the highest order modulation provisioning based on ML-based QoT estimation. The extra information provided by the QoT prediction was used to optimize resource allocation in optical networks [27]. The estimation information is fed into an Integer Linear Programming (ILP) formulation for Routing and Spectrum Assignment (RSA). The results showed that the integration of ML-based QoT estimator and ILP model can improve spectrum efficiency significantly, compared with a traditional ILP-based RSA solution. In addition to the articles which focused on QoT classification or regression tasks, there were some excellent comprehensive review articles published recently. Reference [28] gave a comprehensive perspective on Machine Learning application in optical communication, which mainly focused on the physical layers. The authors also provided a detailed explanation of mathematical foundations for ML algorithms. Reference [29] had an overview of ML explorations in optical communications and networking, the relevant works had been classified and discussed clearly. The introductory tutorial had also been provided for researchers interested in ML.

Currently, most of the works mentioned above focused on QoT estimation of single channel, the relevant algorithms have achieved high accuracy for both classification and regression. For example, most of ML-based regression models for OSNR or Q-factor predictions in [18-21] have achieved the high accuracy for a single channel within at most 1 dB and down to 0.02 dB error, however, the real scenarios of optical systems are generally multi-channel transmission systems. The physical parameters of optical opponents have not been considered in most cases, these physical parameters like EDFA parameters have deep influence on the performance of transmissions.

Therefore, Machine Learning-based QoT estimation should be evaluated over a multi-channel optical network experimentally considering the physical parameters.

III. FIELD-TRIAL TESTBED

Figure 1 shows the 563.4-km field-trial testbed implemented over part of the UK National Dark Fiber Facility (NDFF). The field-trial testbed consists of the NDFF from Bristol to London, the reconfigurable transmitter and transponder nodes in the lab. As shown in Fig. 1, eight external cavity lasers (ECLs) are combined together and then modulated with a dual-polarization IQ modulator to generate 28 Gbaud PM-QPSK signals. The IQ modulator is driven by four 28 Gbaud electrical signals, which are generated by a high-performance FPGA. In a similar way, another eight ECLs are modulated by another IQ modulator to generate 32 Gbaud PM-QPSK signals. **Two sets of PM-QPSK transmitters are used in our setup to generate 16 dummy channels with PM-QPSK signals.** Two Facebook Voyager transponders are used to produce 8-channel 32Gbaud real-time Nyquist PM-16QAM signals as the testing channels for dynamic configuration. The Voyager transponders provide Pre-FEC bit error rates (BERs) for these 8 testing channels. Then, these 24-channel optical signals are multiplexed together by a 4×16 wavelength selective switch (WSS). An optical spectrum analyzer (OSA) measures the optical spectrum of the multiplexed optical signal to assist the WSS for automatic gain equalization. At the receiver side, another 4×16 WSS is used to demultiplex the 8 testing channels for coherent detection. A lab controller is developed to control all the device such as the ECLs, modulators, WSSs, Voyager Transponders and facilities through **GPiB, Ethernet, or USB links**. Finally, the multiplexed optical signals are launched into a 563.4-km loop back field-trial link.

As shown in Fig. 1, the NDFF link consists five nodes with the loop-back configuration at the Telehouse node. Each node includes an 8×8 fiber switch which enables port switching and manages node functions in the architecture-on-demand approach [30]. In the demonstration, each node deployed two gain-constant EDFAs to compensate the link loss for both directions. The length of the link over multiple nodes is **563.4-km**.

A mongo DB-based network configuration and monitoring database (CMDB) is implemented to store network operation information for different testing scenarios. To generate the dataset of various network scenarios, we deploy the node controller and scenario generator to control the state of 8 dynamic channels. As showed in Fig. 2, the 8 testing channels are located in the center while other 16 channels of optical signals act as the neighbor signals. The lab controller configures the lab facility to generate dynamic scenarios with different testing channel switched on. The 16 neighbor channels are switched on all the time, which impact on the 8 testing channels through nonlinear effects in the fibre.

To obtain the dataset, the lab controller will generate different scenarios by configuring the 8-testing channels, including transmitters, the link, and receivers. Each scenario

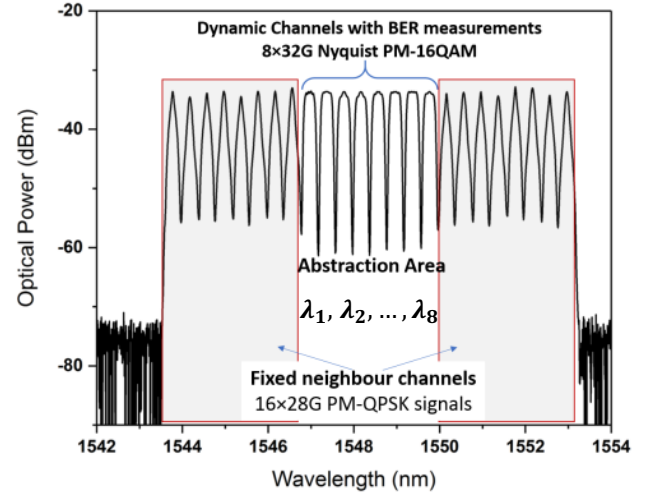


Fig. 2. Optical Spectrum of 24-optical signals configured as ‘ON’ [8].

will be allocated a set of states of the 8-testing channels, i.e., “on” or “off”. The CMDB recorded the monitoring and the performance of the “on” channels linked to each scenario.

Finally, **we prepare the dataset with 1024 samples** which include the network configuration information, monitored optical spectrum, EDFA operation parameters, Q-factors. All the data are collected and stored in CMDB. **Among them, training dataset consists of 899 samples while testing dataset consists of 125 samples. The ratio of training dataset to the testing dataset is 9:1.** ANN is trained on the training dataset to extract the **expected mapping** from the features of network state to the Q-factor converted from the measured BER by Facebook Voyager transponders.

IV. ANN-BASED MULTI-CHANNEL Q-FACTOR PREDICTION

Due to the rapid progress in ML, a lot of ML algorithms have been proposed in academia and industry such as Support Vector Machine, ensemble learning, K-nearest neighbors algorithm, and ANN. Among them, ANN has attracted much more attention recently for its great success in computer vision and natural language processing [31]. **Through the repeated combination of the features and the nonlinear transformation by the activation function, ANN behave as universal function approximators, which can build the expected function from the input to the output through the training.** Besides, ANN has the capability to be scaled up or transferred from one model to another **model** by transfer learning [32]. Several advanced tools are available to build and optimize ANN models such as TensorFlow [33], PyTorch, and Keras [34]. In this work, ANN-based regression model is adopted to estimate multi-channel Q-factors. The proposed ANN-based model includes the following parts.

The first part is the data preparation and feature engineering. With our lab controller, different channels were configured ‘on’ or ‘off’, to generate different scenarios. Then, the network-level CMDB collected network state’s information and real-time Q-factors for the ‘ON’ channels. The network state information

includes optical launcher powers, operation parameters of all the EDFAs. In our experiments, the raw data include over 60 features as link length, launch power, all the EDFA's input power, output power, temperature, gain, laser bias, wavelength information, and the configuration state. **In order to show these features clearly, the TABLE which summarizes the features is provided in appendix. Empirically, the direct training on the original data is likely to cause the over-fitting, because the complexity of ANN generally improves with the increasing of the feature dimensions. The collected dataset is not enough to support the complexed model.** Through feature engineering and intuitive analysis, we observe that the launch power, EDFA's input power, output power, laser bias, and spectrum information have a strong relationship with the Q-factor. In addition to this, a combined vector \vec{V} that contains the configuration information and spectrum information was added into the input features. For example, if channels 1, 2, and 6 are switched on, the state indicator's vector $[1, 1, 0, 0, 0, 1, 0, 0]$ is used to represent the channel state. **Then the relevant wavelengths of 8 testing channels λ_i are combined with this vector as:**

$$\begin{aligned}\vec{V} &= [1, 1, 0, 0, 0, 1, 0, 0] * [\lambda_1, \lambda_2, \lambda_3, \lambda_4, \lambda_5, \lambda_6, \lambda_7, \lambda_8] \\ &= [\lambda_1, \lambda_2, 0, 0, 0, \lambda_6, 0, 0].\end{aligned}\quad (1)$$

This vector can effectively represent the state of each wavelength. The element '0' in the vector means that the corresponding channel is an open spectrum slot. The element '1' indicates the corresponding channel is occupied with optical signal. The corresponding outputs of Q-factor for these channels are recorded in the implemented CMDB. After training, it may build a strong mapping from the channel state vector \vec{V} to the output vector. **Because the vector \vec{V} has strong relationship with the Q-factors of testing channels, the 0 elements in these two 8-dimensional vectors share the same position.**

The second part is to optimise the critical hyperparameters of the ANN model, to achieve high accuracy for the training as well as the testing. ANN has shown the capability to fit an accurate model through training. However, it is not a trivial task to optimize hyperparameters such as the number of hidden layers, the size of each hidden layer, learning rate, batch size, activation function. It will take a long time to manually adjust these hyperparameters, so we use an auto-searching method to optimize the hyperparameters.

We set some hyperparameters of ANN as variables rather than all the hyperparameters. To reduce the required time for the hyperparameters searching, some hyperparameters are initiated empirically. The size of the first layer remains the same as 36 to ensure that the ANN-based model does not lose any important features for the first transformation from the input to the hidden layers. We set the sizes of two hidden layers N^1 , N^2 , regularization index α , the batch size b and the epoch number e as variables with the constraints that $N^1 \in \{20, 24, 28, 32\}$, $N^2 \in \{16, 20, 24\}$, $\alpha \in \{10^{-4}, 10^{-5}, 10^{-6}\}$, $b \in \{5, 10, 15\}$, $e \in \{100, 150, 200\}$. Regarding the activation functions, ReLU function is selected as the activation function for hidden layers. The ReLU function enables not only the nonlinear

transformation effects, but also outputs the non-negative value, which is consistent with the actual value of Q-factors. For the output layers, the Min-Max method has been used to normalize the features and the output. Therefore, the Sigmoid function is more suitable to provide the same range of the normalized Q-factor.

Then, we build all the possible ANN models with the initial hyperparameters and the variable hyperparameters. After searching for the hyperparameters to achieve the minimum MSE on training datasets, the proposed ANN's structure is showed in TABLE I as $[36 \times 24 \times 24 \times 8]$. The sizes of two hidden layers are both 24, **the total number of trainable parameters for ANN is 1668.** The regularization index is 0.000001; The best batch size and epoch number are 10 and 150 respectively.

It should be noted that the mentioned hyperparameters is the optimal results under the constraints we define. The more refined search is needed to determine the global optimal parameters. In fact, we need to tradeoff between the time cost and the accuracy. Compared with the model in [1], we add the Sigmoid function to the output layers to collaborate with the normalization in feature engineering. And we also decrease the regularization index and increase the epochs number for the fine tuning of ANN's weights. The results show that the adjustment indeed improves the stability and generalization of proposed ANN model. **We will give the details of the improvement in the analysis of ANN's learning curve in next section.**

TABLE I
THE SUMMARY OF ANN MODEL

Layer	shape	Activation function	parameters
Input	36	-	-
Hidden 1	24	ReLU	888
Hidden 2	24	ReLU	600
Output	8	Sigmoid	200
Total parameters: 1668; Trainable parameters: 1668			

V. RESULTS AND DISCUSSION

To achieve automatic link setup in a low-margin network, the deployed QoT prediction needs to predict the Q-factors for all the available wavelength slots based on current network status. In addition, the QoT prediction also needs to evaluate the impact of the establishing channel on the current optical channels, considering low-margins are resaved for each channels. With this information, the centralized controller can maximize the link capacity by choosing the best channel with a maximized Q-factor and guarantee the establishing channel does not fail any current channel, or minimize the impacts on the current channels. The trained ANN-based QoT prediction model can predict multiple channels simultaneously. It gives the potential Q-factor for the selected available spectrum slot and also predict the Q-factors for all the current channels when

the selected channel is established.

The performance of proposed ANN-based model for Q-factor regression is presented in this section, the properties of convergence and generalization are also analyzed in detail.

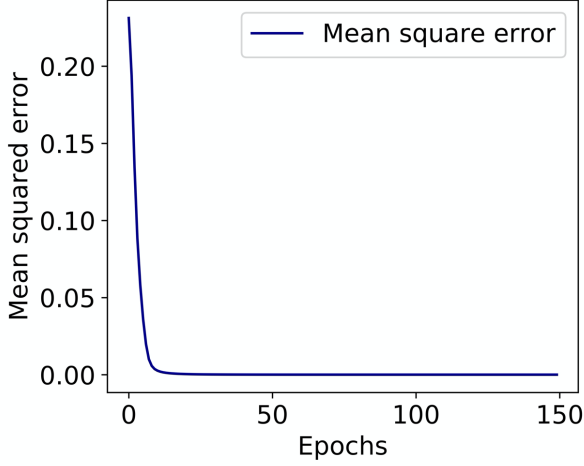


Fig. 3. Mean square error vs. the number of epochs for the training of ANN.

As mentioned above, the proposed ANN is 4 layers including 1668 trainable parameters. The training time is less than 15 seconds on a personal computer (i7-4790, 3.6GHz, 16GB RAM). The predicting time is around several milliseconds. The short training and predicting time make the proposed ANN method suitable for the real-time prediction. Fig. 3 presents the result of MSE vs. epochs in the training process. As the training begins, the MSE drops dramatically, then it decreases slowly with the training. After 35 epochs, the training MSE has reached less than 0.01, and it converges after about 150 epochs.

The learning curve of proposed ANN-based model on training dataset is shown in Fig. 4. The parameter of training size is set from 10% to 100%, the interval is 10%; The seven-fold cross-validation is used to generate the learning curve by dividing all the datasets into seven parts. One of seven parts is selected for the testing and the remaining parts are used for the training, so our model can be tested on seven different datasets.

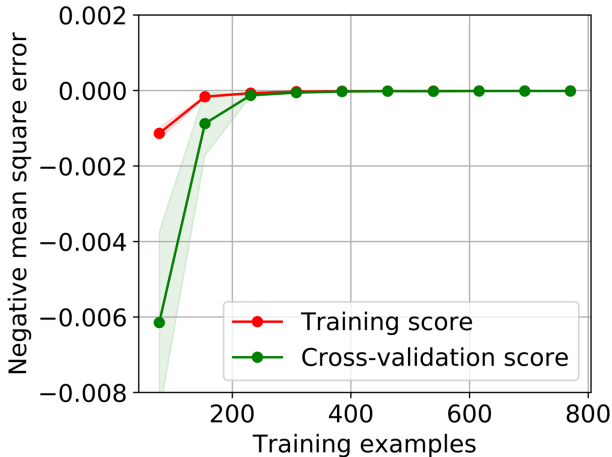


Fig. 4. Learning curve of the implemented ANN model.

We obtain seven negative MSEs for the training and testing of ANN model trained on different datasets. The mean of seven negative MSEs for the training and testing are plotted as the training scores and cross-validation scores. The shadows of two lines are their corresponding standard deviations. As described in Fig. 4, the training scores and cross-validation scores are improved with the increasing of training datasets. When the number of training datasets is more than 250, the mean negative MSE (scores) for both training and cross-validation nearly converges to 0, and the standard deviation of MSEs for the training and cross-validation are very small. The minimum and stable MSEs for cross-validation prove that the proposed ANN model can achieve generalization trained on different datasets. In addition to this, we find that the cross-validation scores of proposed model is better than the model in [1] when the seven-fold cross-validation is performed on parts of training dataset. For 10 percent and 30 percent training dataset, the cross-validation scores are about -0.0067 and -0.00013 while the corresponding two scores in [1] are about -0.114 and -0.0057 . The higher cross-validation scores of proposed model for parts of training dataset mean that the new model achieves better generalization when the size of training dataset is small.

In Fig. 5, the predicted values of testing data that cover the scenarios with not less than 4 available channels are plotted against their actual Q-factors. The figure shows that all the scatters are near the baseline, which proves that the predicted values of Q-factors are very close to the actual Q-factors for these scenarios. After 150 epochs of training on about 900 data, the MSE of Q-factor for the testing data is about 0.0044, and the mean absolute error (MAE) which denotes the gap between the predicted values and the actual values intuitively is about 0.05 dB for the testing data. In order to investigate the influence of the channel state vector \vec{V} on the accuracy of our model, we repeat the training and testing on the dataset without the channel state vector \vec{V} . we find that the MAE will increase to more than 1 dB, which proves that the vector \vec{V} that contains channel states and wavelength information indeed improves the

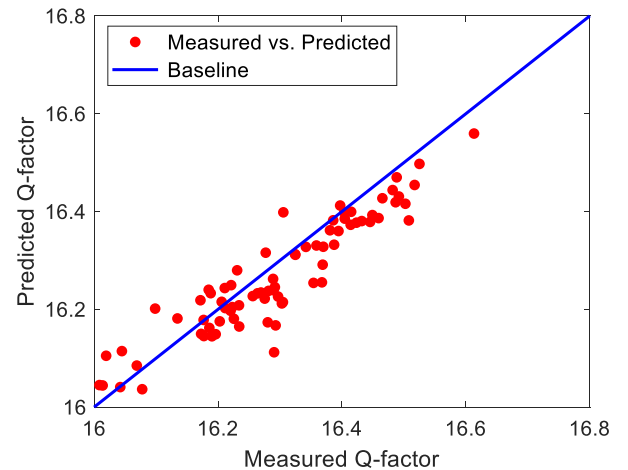


Fig. 5. Predicted Q-factors based on the ANN model vs. measured Q-factors for the scenarios with at least 4 available channels.

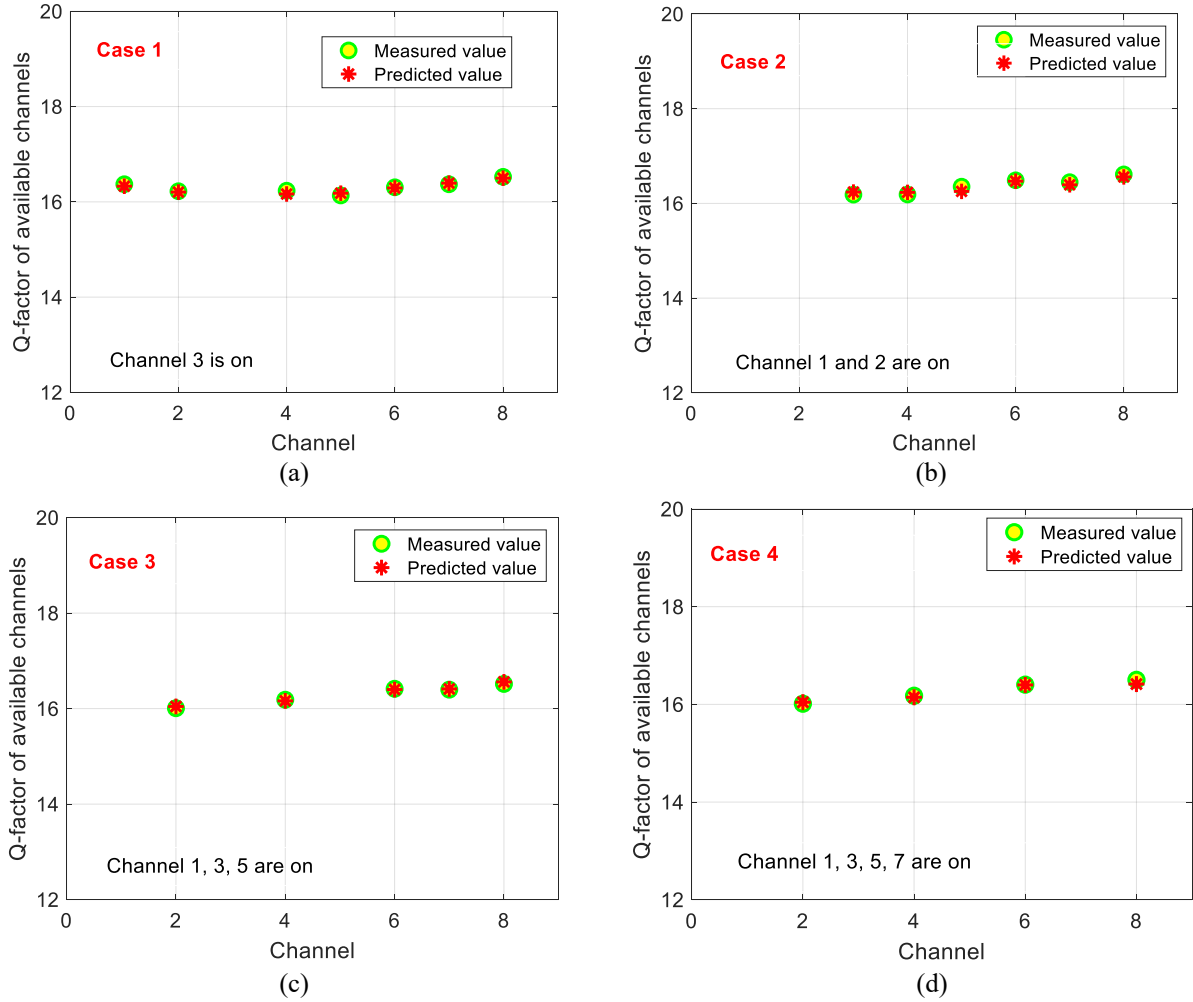


Fig. 6. Predicted Q-factor vs. measured Q-factor for available channels. (a) case 1 for the scenario with channel 3 switched on; (b) case 2 for the scenario with channel 1 and 2 switched on; (c) case 3 for the scenario with channel 1, 3, and 5 switched on; (d) case 4 for the scenario with channel 1, 3, 5, and 7 switched on;

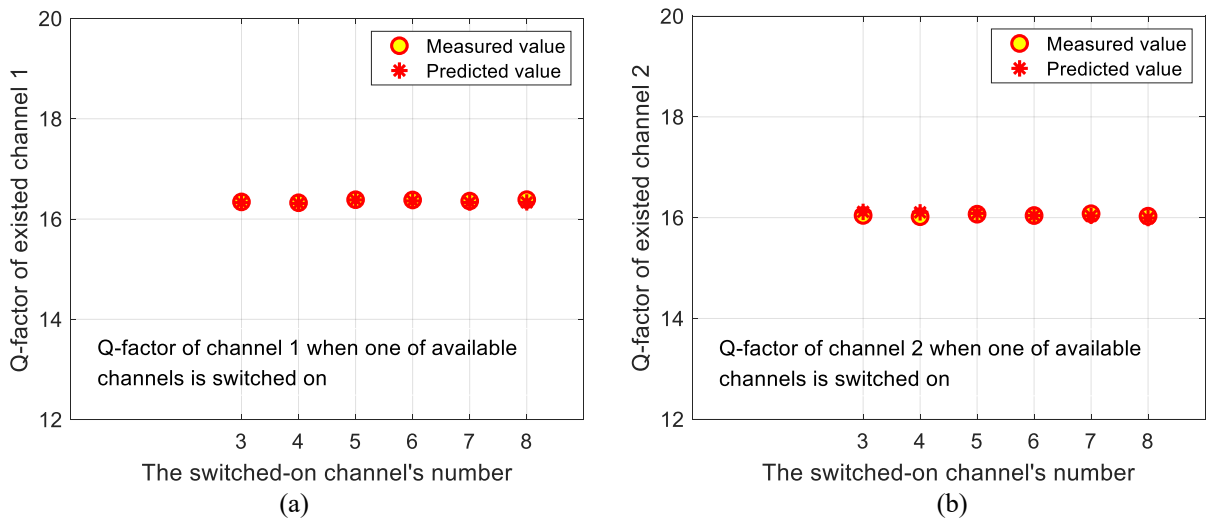


Fig. 7. Predicted Q-factor vs. measured Q-factor for existed channel 1 and 2 when one of available channels is switched on. (a) The prediction for channel 1; (b) The prediction for channel 2.

accuracy of the model.

The performance of ANN-based Q-factor predictor for available channels is evaluated in Fig. 6. We present the

predicted Q-factors of these possible planning channels with the actual measured values. We show the results of four cases in the figure. **In case 1, we remain the channel 3 on and switch on one**

of other 7 channels, then we collect the data for testing. In case 2, we remain channel 1 and 2 on and switch on one of other 6 channels, then we collect the data for testing. In case 3, we keep channels 1, 3, and 5 on and repeat the similar process to collect the data. At last, we keep channels 1, 3, 5, and 7 on for case 4 and collect the data when one of other 4 channels is switched on. The figures for this four scenarios show that the proposed ANN-based Q-factor estimator can make accurate predictions for unestablished channels. The predicted values are very close to the actual values, the model can completely predict the actual Q-factors for available channels. We have optimized the transmission quality for 8 dynamic channels to minimize the mutual interference between different channels. Thus, the actual Q-factors for testing channels are within the range from 16 dB to 16.8 dB. The fact that the Q-factors of two channels are very close requires the ANN-based model to provide QoT predictions with a high accuracy. Therefore, the SDN controller can pick the channel with the highest quality. Fig. 6 show that the accuracy of prediction is satisfactory for SDN controller to distinguish different available channels.

In addition to predicting the Q-factor for new planning channels, we also need to re-evaluate the Q-factor of existed channels to avoid possible interventions from new establishing channels to existed channels. We choose the case 2 in Fig. 6 and re-evaluate the Q-factors of fixed channels 1 and 2 when one of six available channels is switched on. Figure 7 shows that the proposed ANN model can predict existed channel's Q-factor accurately even though the fluctuation of Q-factor for channel 1 and 2 is small. The accurate re-evaluation by ANN-based model can help SDN controller to avoid severe interference when establishing new channels.

The learning curve in Fig. 4 shows that the MSE of proposed ANN model is very small for both training and cross-validation when the size of datasets is more than 250. To investigate the influence of training data's size on ANN model's accuracy, the MSE and MAE of proposed ANN model for all the testing data are presented with different size of training datasets in Fig. 8. We keep the testing dataset unchanged and change the size of training dataset. The figure shows that both MSE and MAE decrease with the size of datasets increasing. The MAE of ANN model has achieved the small value as about 0.11 for 300 training data, and it can converge to about 0.05 for 900 training data. The results proved that the size of training datasets still has a significant influence on ANN to achieve high accuracy.

Finally, some inspirations about the model selection is presented. As mentioned above, the Q-factors of testing channel is within the small range. The distribution of Q-factors for 8 testing channels seems to be linear in some cases, which implies that the linear regression model may be effective to fit the data. The main reason we choose ANN-based model is that ANN can behave as universal function estimators theoretically from the linear model to the complexed nonlinear model through different shape and hyperparameters of ANN. The performance of ANN by auto-search method is often better than hand-tuned models. However, it spends several hours for the auto-search method to select the hyperparameters of ANN if we use a personal PC with limited computing resource. Therefore, the

time cost should be considered when we use auto-search method.

VI. FUTURE WORK

The experimental demonstration shows that the ANN-based Q-factor predictor provides multiple channel information simultaneously based on the collected network operation data from both the SDN controller and the optical performance monitoring. In the current demonstration, the Q-factors of the multiple paths that connect several optical nodes can be predicted precisely with the implemented CMDB. And the developed QoT predictor shows a good generality for all channel configurations (Here we referred to the 8 channels): The ANN-based model will output the Q-factors for the switched-on channels and 0 for the switched-off channels respectively. The work provides a possible way to adopt ANN in optical networks. However, there are still many challenges this paper hasn't addressed. Firstly, complex scenarios such as multiple modulation formats with flexible spectrum allocation have not been investigated. Secondly, the scale up of the current ANN model to a network with complex topologies still needs lots of studies. In the future work, we will further investigate these challenges.

In addition to the single domain prediction, there are more research interests for multi-domain QoT predictions. Distributed training for several independent models is not preferred in terms of the time and computing resource. The efficient retraining of existed ANN models should be considered. Meta Learning-based fast training will be investigated over multi-domain optical networks in the future work.

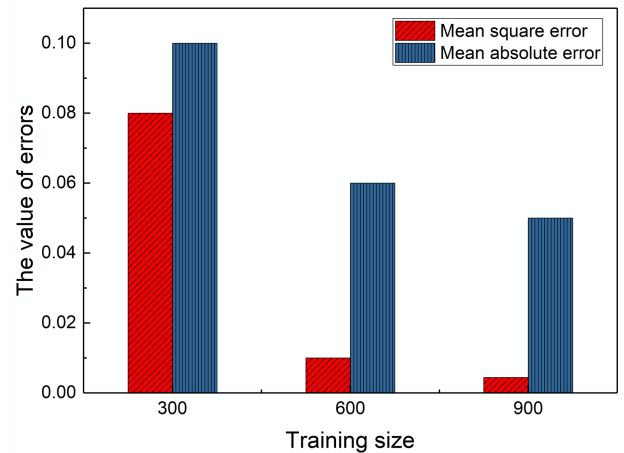


Fig. 8. MSE and MAE of the testing data for ANN model with different size of training dataset.

VII. CONCLUSION

In this paper, ANN-based multi-channel Q-factor prediction is investigated over a 563.4-km field-trial testbed. An auto-search method is used to find the optimized hyperparameters of the ANN model. The results show that the implemented ANN-

TABLE II
THE RAW FEATURES OF SAMPLES IN THE DATASET

Link information	EDFA operation parameters	Configuration state	Receiver
Length1	BRD4 : input power, output power, laser bias, gain, temperature	Channel1 state	Q-factor1
Length2	FFD2 : input power, output power, laser bias, gain, temperature	Channel2 state	Q-factor2
Length3	RDG8 : input power, output power, laser bias, gain, temperature	Channel3 state	Q-factor3
Length4	PGT4 : input power, output power, laser bias, gain, temperature	Channel4 state	Q-factor4
Length5	THN2 : input power, output power, laser bias, gain, temperature	Channel5 state	Q-factor5
Length6	PGT2 : input power, output power, laser bias, gain, temperature	Channel6 state	Q-factor6
Length7	RDG6 : input power, output power, laser bias, gain, temperature	Channel7 state	Q-factor7
Length8	FFD4 : input power, output power, laser bias, gain, temperature	Channel8 state	Q-factor8
Length9	BRD2 : input power, output power, laser bias, gain, temperature	Launch power	
Length10			

based multi-channel estimator converges quickly and achieves generalization for new data. The proposed ANN-based regression model can predict the Q-factor of both new planning channels and existed channels accurately. The precise prediction for new planning channels can assist SDN controller to pick the channel with high transmission quality. The re-evaluation for existed channels can avoid the potentially severe interference from new establishing channels. The comprehensive abstraction information from ANN-based multi-channel Q-factor estimator can help the SDN controller to make a proper decision during the real-time network configuration.

APPENDIX

The raw features of samples in the dataset is presented in TABLE II. In order to observe these features more intuitively, the lengths of spans and all the EDFAs are highlighted in Fig. 9. The link information includes the lengths of all the spans in the testbed from length1 to length10. We see from the figure that the whole link includes 9 EDFAs as BRD4, FFD2, RDG8 etc., we collect each EDFA's input power, output power, laser bias, gain and temperature as the EDFA's operation parameters. For the configuration information, we record the 8 testing channel's state and the launch power. As mentioned above, the channel state is presented as an 8-dimensional vector. The corresponding elements for the switched-on channels is presented as their wavelength. The other elements in this vector are 0. The number of features for the raw dataset is over 60. After the feature engineering and analysis, we select the launch power (1 element), all the EDFA's input power, output power and laser bias (27 elements), and the 8-dimensional vector for the channel state as the final features (totally 36 elements) for the training and testing.

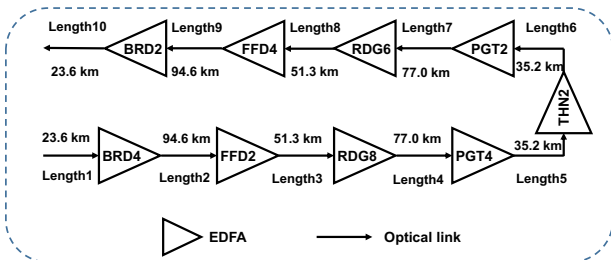


Fig. 9. The structure of links in NDFF.

REFERENCES

- [1] Z. Gao, S. Yan, J. Zhang, M. Mascarenhas, R. Nejabati, Y. Ji, and D. Simeonidou, "ANN-Based Multi-Channel QoT-Prediction over a 563.4-km Field-Trial Testbed," 2019 European Conference on Optical Communication (ECOC), Dublin, Ireland.
- [2] M. Fiorani et al., "Challenges for 5G transport networks," 2014 IEEE International Conference on Advanced Networks and Telecommunications Systems (ANTS), New Delhi, 2014, pp. 1-6.
- [3] Y. Ji, J. Zhang, X. Wang, H. Yu, "Towards converged, collaborative and co-automatic (3C) optical networks," *SCIENCE CHINA INFORMATION SCIENCES*, vol.61, no.12, 121301, 2018.
- [4] Y. Ji, J. Zhang, Y. Xiao and Z. Liu, "5G flexible optical transport networks with large-capacity, low-latency and high-efficiency," in *China Communications*, vol. 16, no. 5, pp. 19-32, May 2019.
- [5] 3GPP TS 22.261.: Service requirements for the 5G system, Stage 1, V16.6.0, Rel.15 (2018).
- [6] J. Ordóñez-Lucena, P. Ameigeiras, D. Lopez, J. J. Ramos-Munoz, J. Lorca and J. Folgueira, "Network Slicing for 5G with SDN/NFV: Concepts, Architectures, and Challenges," *IEEE Communications Magazine*, vol. 55, no. 5, pp. 80-87, May 2017.
- [7] J. Zhang, Y. Ji, S. Jia, H. Li, X. Yu and X. Wang, "Reconfigurable optical mobile fronthaul networks for coordinated multipoint transmission and reception in 5G," in *IEEE/OSA Journal of Optical Communications and Networking*, vol. 9, no. 6, pp. 489-497, June 2017.
- [8] Yan, S., Gao, Z., Wang, R., et al.: 'Dynamic Abstraction of Optical Networks with Machine Learning Technologies', ONDM., Greece., Mar. 2019, pp. 1-6.
- [9] D. Rafique and L. Velasco, "Machine learning for network automation: overview, architecture, and applications [Invited Tutorial]," in *IEEE/OSA Journal of Optical Communications and Networking*, vol. 10, no. 10, pp. D126-D143, Oct. 2018.
- [10] Y. Pointurier, "Design of low-margin optical networks," in *IEEE/OSA Journal of Optical Communications and Networking*, vol. 9, no. 1, pp. A9-A17, Jan. 2017.
- [11] P. Poggiolini, "The GN Model of Non-Linear Propagation in Uncompensated Coherent Optical Systems," *Journal of Lightwave Technology*, vol. 30, no. 24, pp. 3857-3879, Dec.15, 2012.
- [12] J. Shao, X. Liang and S. Kumar, "Comparison of Split-Step Fourier Schemes for Simulating Fiber Optic Communication Systems," *IEEE Photonics Journal*, vol. 6, no. 4, pp. 1-15, Aug. 2014.
- [13] P. Samadi, D. Amar, C. Lepers, M. Lourdiane and K. Bergman, "Quality of Transmission Prediction with Machine Learning for Dynamic Operation of Optical WDM Networks," 2017 European Conference on Optical Communication (ECOC), Gothenburg, 2017, pp. 1-3.
- [14] C. Rottondi, L. Barletta, A. Giusti and M. Tornatore, "Machine-learning method for quality of transmission prediction of unestablished lightpaths," in *IEEE/OSA Journal of Optical Communications and Networking*, vol. 10, no. 2, pp. A286-A297, Feb. 2018.
- [15] S. Aladin and C. Tremblay, "Cognitive Tool for Estimating the QoT of New Lightpaths," 2018 Optical Fiber Communications Conference and Exposition (OFC), San Diego, CA, 2018, pp. 1-3.
- [16] R. M. Morais and J. Pedro, "Evaluating Machine Learning Models for QoT Estimation," 2018 20th International Conference on Transparent Optical Networks (ICTON), Bucharest, 2018, pp. 1-4.

- [17] T. Panayiotou, G. Savva, B. Shariati, I. Tomkos and G. Ellinas, "Machine Learning for QoT Estimation of Unseen Optical Network States," 2019 Optical Fiber Communications Conference and Exhibition (OFC), San Diego, CA, USA, 2019, pp. 1-3.
- [18] S. Yan et al., "Field Trial of Machine-Learning-Assisted and SDN-Based Optical Network Management," 2019 Optical Fiber Communications Conference and Exhibition (OFC), San Diego, CA, USA, 2019, pp. 1-3.
- [19] R. Proietti et al., "Experimental demonstration of machine-learning-aided QoT estimation in multi-domain elastic optical networks with alien wavelengths," *IEEE/OSA Journal of Optical Communications and Networking*, vol. 11, no. 1, pp. A1-A10, Jan. 2019.
- [20] E. Seve, J. Pesic, C. Delezoide, S. Bigo and Y. Pointurier, "Learning process for reducing uncertainties on network parameters and design margins," *IEEE/OSA Journal of Optical Communications and Networking*, vol. 10, no. 2, pp. A298-A306, Feb. 2018.
- [21] I. Sartzetakis, K. K. Christodouloupoulos and E. M. Varvarigos, "Accurate quality of transmission estimation with machine learning," *IEEE/OSA Journal of Optical Communications and Networking*, vol. 11, no. 3, pp. 140-150, March 2019.
- [22] J. Yu, W. Mo, Y. Huang, E. Ip and D. C. Kilper, "Model transfer of QoT prediction in optical networks based on artificial neural networks," *IEEE/OSA Journal of Optical Communications and Networking*, vol. 11, no. 10, pp. C48-C57, October 2019.
- [23] Le Xia, Jing Zhang, Shaohua Hu, Mingyue Zhu, Yingxiong Song, and Kun Qiu, "Transfer learning assisted deep neural network for OSNR estimation," *Opt. Express*, vol. 27, no. 14, pp. 19398-19406, 2019.
- [24] D. Azzimonti, C. Rottondi and M. Tornatore, "Reducing probes for quality of transmission estimation in optical networks with active learning," in *IEEE/OSA Journal of Optical Communications and Networking*, vol. 12, no. 1, pp. A38-A48, January 2020.
- [25] S. Yan et al., "Field trial of Machine-Learning-assisted and SDN-based Optical Network Planning with Network-Scale Monitoring Database," 2017 European Conference on Optical Communication (ECOC), Gothenburg, 2017, pp. 1-3.
- [26] S. Oda, M. Bouda, O. Vassiliev, Y. Hirose, T. Hoshida and T. Ikeuchi, "Network Capacity Improvement by Quality of Transmission Estimator with Learning Process," 2017 European Conference on Optical Communication (ECOC), Gothenburg, 2017, pp. 1-3.
- [27] M. Salani, C. Rottondi and M. Tornatore, "Routing and Spectrum Assignment Integrating Machine-Learning-Based QoT Estimation in Elastic Optical Networks," IEEE INFOCOM 2019 - IEEE Conference on Computer Communications, Paris, France, 2019, pp. 1738-1746.
- [28] F. N. Khan, Q. Fan, C. Lu and A. P. T. Lau, "An Optical Communication's Perspective on Machine Learning and Its Applications," *Journal of Lightwave Technology*, vol. 37, no. 2, pp. 493-516, Jan. 2019.
- [29] F. Musumeci et al., "An Overview on Application of Machine Learning Techniques in Optical Networks," *IEEE Communications Surveys & Tutorials*, vol. 21, no. 2, pp. 1383-1408, Secondquarter 2019.
- [30] N. Amaya, G. Zervas, and D. Simeonidou, "Introducing node architecture flexibility for elastic optical networks," *IEEEOSA J. Opt. Commun. Netw.*, vol. 5, no. 6, pp. 593-608, 2013.
- [31] LeCun, Y., Bengio, Y. & Hinton, G. Deep learning. *Nature* 521, pp. 436-444, 2015.
- [32] S. J. Pan and Q. Yang, "A Survey on Transfer Learning," in *IEEE Transactions on Knowledge and Data Engineering*, vol. 22, no. 10, pp. 1345-1359, Oct. 2010.
- [33] Abadi et al., "Tensorflow: A system for large-scale machine learning," Tech. rep., Google Brain, 2016. arXiv preprint.
- [34] Ketkar N. Introduction to PyTorch in Deep Learning with Python. Apress, Berkeley, CA, 2017.